RadLoco: A Rapid and Low Cost Indoor Location-Sensing System

Eugene Hyun, Michael McGuire, Mihai Sima

Department of Electrical and Computer Engineering
University of Victoria

Chinacom 2008 Conference
Outline

1 Location Aware Computing
   - Motivation
   - Current Technologies
   - Problem Definition

2 Location Sensing System
   - Overview
   - Data Collection
   - Location Estimation
   - Simple Graphical Example

3 Experiment and Results
   - ECS Room 116
   - ECS 5th and 6th floors

4 Conclusions
Motivation

Location aware computing consists of the use of location information to improve the value of a wireless network for the users.

Examples

- providing navigation through unfamiliar environments
- dynamic pre-allocation of resources
**Motivation**

Location aware computing consists of the use of location information to improve the value of a wireless network for the users.

**Examples**

- providing **navigation** through unfamiliar environments
- dynamic **pre-allocation** of resources
Location Aware Computing

Motivation

Location aware computing consists of the use of location information to improve the value of a wireless network for the users.

Examples

- providing **navigation** through unfamiliar environments
- dynamic **pre-allocation** of resources
Current Technologies for Location-Sensing

Global Position System (GPS) and Cellular Networks
- poor location performance indoors

Wireless Local Area Networks
- Many environments have ubiquitous wireless networks
- Wireless network access is being incorporated into smaller mobile devices
Current Technologies for Location-Sensing

Global Position System (GPS) and Cellular Networks
- Poor location performance indoors

Wireless Local Area Networks
- Many environments have *ubiquitous* wireless networks
- Wireless network access is being incorporated into smaller mobile devices
Current Technologies for Location-Sensing

Global Position System (GPS) and Cellular Networks
- poor location performance indoors

Wireless Local Area Networks
- Many environments have ubiquitous wireless networks
- Wireless network access is being incorporated into smaller mobile devices
Problem Definition

**Problem**
- The relationship between location and radio signal strength is highly **non-linear** and not known a priori.
- Survey data collection is **time consuming**.

**Solution**
- Use non-parametric **estimation technique** to reduce noise and required number of survey points.
- **Sensory network** of Ultrasonic/Radio devices to aid in rapid survey data collection.
Problem Definition

**Problem**
- The relationship between location and radio signal strength is highly **non-linear** and not known a priori.
- Survey data collection is **time consuming**.

**Solution**
- Use non-parametric **estimation technique** to reduce noise and required number of survey points.
- **Sensory network** of Ultrasonic/Radio devices to aid in rapid survey data collection.
Location Sensing System Overview

Data Collection

Cricket Sensor Network
Ground Truth Location \((x, y)\)

WLAN
WLAN data = (ssid, MAC, RSS)

Data Processing

Kernel/Parzen Estimator

Estimated Location

Details

2 stages: Data Collection and Data Processing

Eugene Hyun, Michael McGuire, Mihai Sima

RadLoco: A Rapid and Low Cost Indoor Location-Sensing System
Location Sensing System Overview

**Data Collection**
- Cricket Sensor Network
  - Ground Truth Location (x, y)

**Data Processing**
- WLAN
  - WLAN data = (ssid, MAC, RSS)
- Kernel/Parzen Estimator
- Estimated Location

**Details**

Data Collection: Cricket Sensory Network
Cricket Sensory Network for Ground Truth Location

**Sensory Network**
- provide **ground truth location** for survey point.
- use **modified steepest descent** and Newton’s method optimization

**Cricket Details**
- **accuracy** = mean accuracy of 14 cm
- **sensors** = ultrasonic (40 kHz), radio (433 MHz)
- **developers** = Networks and Mobile Systems Group at MIT.
- **cost** = $3000 for 12 crickets covering 500 m²
Cricket Sensory Network for Ground Truth Location

Sensory Network
- provide **ground truth location** for survey point.
- use **modified steepest descent** and Newton’s method optimization

Cricket Details
- **accuracy** = mean accuracy of 14 cm
- **sensors** = ultrasonic (40 kHz), radio (433 MHz)
- **developers** = Networks and Mobile Systems Group at MIT.
- **cost** = $3000 for 12 crickets covering 500 m²
Location Sensing System Overview

Data Collection
- Cricket Sensor Network
  Ground Truth Location (x, y)
- WLAN
  WLAN data = (ssid, MAC, RSS)

Data Processing
- Kernel/Parzen Estimator
- Estimated Location

Details
Data Collection: WLAN

Eugene Hyun, Michael McGuire, Mihai Sima
RadLoco: A Rapid and Low Cost Indoor Location-Sensing System
WLAN information

WLAN data

- **RSS** from base station
- **MAC** of base station
- Time
- SSID (network ID)

Survey Data

- WLAN data and ground truth location compose of a **survey point**.
- Collect several survey points within a floor or room to create a **survey data set**.
WLAN information

**WLAN data**
- **RSS** from base station
- **MAC** of base station
- Time
- SSID (network ID)

**Survey Data**
- WLAN data and ground truth location compose of a **survey point**.
- Collect several survey points within a floor or room to create a **survey data set**.
Location Sensing System Overview

**Location Sensing System**

Data Collection

- Cricket Sensor Network
  Ground Truth Location \((x, y)\)

WLAN

WLAN data = (ssid, MAC, RSS)

Data Processing

- Kernel/Parzen Estimator

Estimated Location

Details

Data Processing: Location Estimation

Eugene Hyun, Michael McGuire, Mihai Sima

RadLoco: A Rapid and Low Cost Indoor Location-Sensing System
Location Estimation

Estimation Algorithm

Non-Parametric **Kernel/Parzen Estimation** Technique using a sum of multivariate Gaussian Distributions.

Variables

- $b = \text{vector of base stations (WLAN)}$
- $z = \text{vector of receive signal strength measurements (WLAN)}$
- $\theta = x \& y \text{ coordinates (survey points)}$
- $\hat{\theta} = x \& y \text{ coordinate (estimated location)}$
Location Estimation

Estimation Algorithm

Non-Parametric **Kernel/Parzen Estimation**
Technique using a sum of multivariate Gaussian Distributions.

Variables

- $b =$ vector of base stations (WLAN)
- $z =$ vector of receive signal strength measurements (WLAN)
- $\theta =$ x & y coordinates (survey points)
- $\hat{\theta} =$ x & y coordinate (estimated location)
Simple Graphical Example

3 base stations and 4 cricket sensors
Simple Graphical Example

Explanation

crickets provide ground truth location (blue dots)
RadLoco: A Rapid and Low Cost Indoor Location-Sensing System

Simple Graphical Example

Explanation

base stations provide RSS vector (z)
Simple Graphical Example

Explanation

take RSS measurement at current unknown location
**Simple Graphical Example**

Explanation

and use the survey points collected prior

**Overview**

Data Collection

Location Estimation

Conclusions
Simple Graphical Example

Explanation
to calculate the weight from \( z \) feature and Kernel estimator.
Simple Graphical Example

Explanation

estimated location is calculated using weights and $\theta$ feature
ECS 116 - Experimental Setup

Environment

- Engineering and Computer Science building lecture hall 116
- 117 seats divided into 8 rows
- 10 x 13 meters with 2 seats/m

Equipment

- hardware: Dell Inspiron laptop + Intel 802.11g WLAN card
- software: Netstumber, Radloco, MySQL
# ECS 116 - Experimental Setup

## Environment
- Engineering and Computer Science building lecture hall 116
- 117 seats divided into 8 rows
- 10 x 13 meters with 2 seats/m²

## Equipment
- hardware: Dell Inspiron laptop + Intel 802.11g WLAN card
- software: Netstumber, Radloco, MySQL
ECS 116 - Experimental Results

Description

- We were able to obtain over 80% accuracy within 4 meters or 6 seats.
- Motivation for ECS 5th and 6th floor
ECS 5th and 6th floors - Experimental Setup

**Environment**
- Building has an *atrium* like structure
- Experiments were restricted to *hallways* and public conference rooms.

**Base Station (WAP) Information**
- WAPs from networks *UVicOpen* or *engrnet*.
- More than 35 unique WAPs and minimum of 12 WAPs visible at any location on a floor.
ECS 5th and 6th floors - Experimental Setup

Environment
- Building has an atrium like structure
- Experiments were restricted to hallways and public conference rooms.

Base Station (WAP) Information
- WAPs from networks UVicOpen or engrnet.
- more than 35 unique WAPs and minimum of 12 WAPs visible at any location on a floor.
Experimental Results - Decimation of Survey Set

Description

Decimating survey set yields near identical performance.
Experimental Results - 6th vs. 5th floor

Description
Both floor have similar accuracy of **3.5 m** more than 80% of the time.
Experimental Results - Number of Base Stations

Description

Increasing base stations yields better accuracy

Eugene Hyun, Michael McGuire, Mihai Sima

RadLoco: A Rapid and Low Cost Indoor Location-Sensing System
Experimental Results - Combined Datasets

Description

Combined floors data sets provides 3D location accuracy with over **98% floor accuracy**
Major Contributions

- location-sensing system that locates mobile computing devices indoors based on **WLAN** technology.
- **sensory network** composed of radio/ultrasonic devices allow rapid data collection
- **98% accuracy** floor accuracy and location estimations within **3.5m** of true location
Conclusions

Further Research

- IEEE 802.16, WiMAX, can be extended for indoor location.
- Sensory network **calibration** on walls
- **real-time** location estimation for the RadLoco software
### List of Publications


Thank you!

Eugene Hyun, Michael McGuire, Mihai Sima

RadLoco: A Rapid and Low Cost Indoor Location-Sensing System
Block Diagram of Estimated Location

weighted average function

\[ w_i = \frac{K(x - X_i)}{\sum_{i=1}^{N} K(x - X_i)} \]  (1)
**Software Setup**

**Steps**

Start **Netstumbler** and load Radloco perlscript

---

**Eugene Hyun, Michael McGuire, Mihai Sima**

**RadLoco: A Rapid and Low Cost Indoor Location-Sensing System**
Software Setup

steps
load map into Radloco and click on map to record data
Software Setup

steps

- Crickets provide ground truth coordinates
Software Setup

**steps**

RSS, BSSID and coordinates are saved to MySQL database

**RSS**, **BSSID** and **coordinates** are saved to MySQL database
Software Setup

**steps**
data processing is performed in Matlab
The MMSE of the terminal location is

$$\hat{\theta}_{\text{MMSE}} = E[\theta|x] = \int_S \theta f_\Theta(\theta|x) d\theta \quad (2)$$

The above equation can be expanded to

$$\hat{\theta}_{\text{MMSE}} = \int_S \theta \frac{f_{X,\Theta}(x,\theta)}{f_X(x)} d\theta = \frac{\int_S \theta f_{X,\Theta}(x,\theta) d\theta}{\int_S f_{X,\Theta}(x,\theta) d\theta} = \frac{\int_S \theta f_{X,\Theta}(x,\theta) d\theta}{\int_S f_X(x) d\theta} \quad (3)$$
Bayesian Estimation Algorithm - 2

Minimum Mean Square Estimate (MMSE)

The joint approximate PDF is given by

$$\hat{f}_{\Theta, x}(\theta, x) = \frac{1}{N} (h_x)^{-L} (h_\theta)^{-2} \sum_{i=1}^{N} K_x \left( \frac{x - X_i}{h_x} \right) K_\theta \left( \frac{\theta - \theta_i}{h_\theta} \right)$$  \hspace{1cm} (4)

with $K(\cdot)$ being the Kernel functions for location and RSS measurements. The value $L$ represents length of vector $x$.

Kernel Function

For the kernel functions $K(\cdot)$ we use the standard multivariate Gaussian distribution:

$$K_x(x) = \left( \frac{1}{\sqrt{2\pi}} \right)^L \exp \left( -\frac{1}{2} (x^T x) \right)$$  \hspace{1cm} (5)
If we substitute the estimated PDF from Equation (4) into (3) and perform integration by substitution using $u = \theta - \theta_i$, we can reduce:

$$\int \theta K_\Theta (\theta - \theta_i) d\theta = \int (u + \theta_i) K_U(u) du = \theta_i$$

since the mean of the random variable $U$ is:

$$\int u K_U(u) du = 0$$

and

$$\int K_U du = 1$$
The resulting estimated location $\hat{\theta}$ is then a weighted average where the sum is taken across the weighted locations in the subset:

$$\hat{\theta} = \sum_{i=1}^{N} w_i \theta_i$$  \hspace{1cm} (7)

with the weights $w_i$ for each survey point $i$ being defined as:

$$w_i = \frac{K(x - X_i)}{\sum_{i=1}^{N} K(x - X_i)}$$  \hspace{1cm} (8)
Kernel width

Small values of $h_x$ indicate that the RSS vectors can change radically with short spatial displacements of the mobile device while larger values indicate that the RSS changes significantly only with large spatial displacements.
two datasets are needed for training:

1. survey set - dataset A
2. validation set - dataset B.

Each location in dataset B estimated using Kernel Estimator with survey dataset A, while varying the parameter $h_x$.

The $h_x$ that produces the minimum MSE is the optimal kernel width for the training dataset B.

In theory, we need a training set of infinite size to determine optimal kernel. Therefore optimal means best kernel given our finite datasets.

A third dataset C can be used for verification and is considered the test dataset. The set C is not used for training because it would extend the survey set.
collecting multiple datasets may be costly and time consuming: determine the optimal kernel width $h_x$ for one dataset using a cross-validation approach.

$$MSE(h) = \sum_{i=1}^{N} \left[ \theta_i - \sum_{k=1, \neq i}^{N} w_k \theta_k \right]^T \times \left[ \theta_i - \sum_{k=1, \neq i}^{N} w_k \theta_k \right]$$ (9)

remove a survey point from the survey set, then estimating the location of the survey point using the rest of the survey points, while varying $h_x$ values.
Error Calculation

Explanation

*green dot* = true location
*red dot* = estimated location
Error Calculation

Explanation

error distance = \sqrt{(p_x - \hat{p}_x)^2 + (p_y - \hat{p}_y)^2)}
error distance = \sqrt{(p_x - \hat{p}_x)^2 + (p_y - \hat{p}_y)^2)